network selectively couples the local area networks to each other through links created between its provider edge routers. To support operation, the provider edge routers typically maintain Virtual Routing and Forwarding (VRF) information dictating how to route and forward traffic through the shared physical network to support corresponding VPNs for the different customers.

According to one conventional technique, a service network may be extended beyond provider edge nodes to customer edge nodes. For example, the connectivity model supported by RFC2547 generally enables any CE (Customer Edge) nodes to establish a link between each other for transmission of data messages between corresponding interconnected networks.

SUMMARY

Unfortunately, there are deficiencies associated with conventional techniques of configuring edge routers with configuration data identifying how to forward data packets across an associated service provider network. For example, as discussed above, [Request For Comment 2547] supports a service architecture capable of building and supporting any data path between CE routers. However, such a model does not provide any inherent data encryption services. Therefore, customers that wish to encrypt their traffic must do so before it enters the [RFC2547] network. Although burdensome, this is typically achieved by enabling IPSec (Internet Protocol Security) encryption and running IPSec tunnels between CE routers that belong to a corresponding "encrypted" VPN.

Some service providers have provided at least partial integration of [RFC2547] and IPSec by terminating IPSec tunnels into a Virtual Routing & Forwarding Instance (VRF) at PE (Provider Edge) routers; however, this provides a secure path between the PE and CE devices but does not provide an end-to-end security between the CE nodes. Deployment of IPSec tunnel meshes is analogous to the "overlay" model used in Frame-relay or ATM (Asynchronous Transfer Mode) networks. Although network-based, this conventional solution of terminating IPsec tunnels at PE routers does not extend IPSec between security gateways of different customer sites.

It is an advancement in the art to combine the CE-to-CE protection methodologies with the "any-to-any" node connectivity model provided by [RFC2547] so that the customer experience is seen as the "best of both worlds." Accordingly, one embodiment of the present invention includes an apparatus and method for disseminating configuration data such as routing policy information to edge nodes of a network supporting virtual networks without, e.g., a need for using direct routing protocol exchange or IP based tunnels such as those provided by GRE (Generic Routing Encapsulation). This enhanced model supports CE-to-CE data protection while eliminating the requirement for pre-established IP-based tunnels or routing adjacencies between [IPSec] security gateways.

More specifically, an embodiment of the present invention involves generating a notification message (at a first node of a physical network) to include routing policy attributes such as network address information and a corresponding gateway identifier. The gateway identifier identifies a gateway in the physical network through which future generated data messages shall be forwarded to at least one host computer (e.g., any computer having an associated network address) as indicated by the network address information of the gateway identifier. In other words, the network address information identifies a single or range of network addresses of computers.

The first node transmits or distributes the notification message to a second node (or multiple nodes) of the physical network, thereby enabling the second node (and potentially other nodes) to establish a secure virtual network connection between the second node (or other nodes) and the first node or other node specified by the gateway identifier on which to forward data messages to the one or more host computer(s) based on the corresponding gateway identifier. For example, the second node may generate and maintain a map including configuration data (based on the routing policy attributes such as the network address information and gateway identifier) so that data traffic through the second node to the at least one computer identified by the network address information is forwarded through the gateway as identified by the gateway identifier. Based on this technique of disseminating routing policy attributes, nodes (e.g., CE routers) of a network can be dynamically configured to support routing of messages based on receipt of the

network address information and gateway identifier (e.g., an IPsec identity such as Ipv4, distinguished name, etc.).

A corresponding map at the second node, after being dynamically updated based on the routing policy attribute, includes configuration data indicating on which virtual network to forward data depending on the address (or subnet) associated with a host computer to which the data messages are forwarded. For example, as previously mentioned, the notification message includes a gateway identifier and network address information. Configuration data associated with the virtual network stored in the map may identify a VPN connection (including the gateway as identified by the gateway identifier) on which to forward data messages destined for computers identified by the corresponding network address information.

Embodiments of the invention thus allow a network administrator of a customer network at a first facility (e.g., a first LAN location for the customer) that operates on a service provider network to automatically disseminate (e.g., using a protocol such as the Border Gateway Protocol), in a notification message, security gateway information and corresponding network address information to other customer facility locations (e.g., other LANs for that same customer located elsewhere on the service provider network). The network address information identifies a set of host computer systems (e.g., an address range or list of one or more computers) within that first LAN facility (e.g., a IP subnet address range, or a list of host IP address) that must be communicated with, from other customer LAN facilities, in a secure manner (e.g., using a VPN). The corresponding security gateway identifier indicates the endpoint or termination point of that secure connection through which communications must pass in order to communicate with those hosts that have addresses that match the network address information (i.e., hosts within the specified address range). As an example, if a subnet in the LAN facility contains a group of sensitive computers that must be communicated with (from hosts outside of that LAN facility) in a secure manner, the notification message can contain network address information identifying the subnet of all of those hosts. The gateway identifier can contain the address of the firewall or other gateway device, such as the customer edge router that couples that LAN facility to the service

provider network. That gateway device (i.e., the customer edge router) is pre-configured with VPN software or hardware to accept secure VPN connections. As such, when embodiments of the invention disseminate this notification information to remotely located customer edge routers at other customer facilities (i.e., that link other LANs for that customer to the service provider network), those other customer edge routers can use this information to establish a secure connection to the device having the IP address of the gateway identifier when any communications (e.g., packets) are destined for a host computer that falls within the range of address(es) specified by the network address information.

Upon receipt of a data message at the second node such as a data packet to be transmitted over a service provider network, the second node identifies a destination address of the data message. If the destination address (and potentially also a source address) of the data message matches an entry in the corresponding (routing) map at the second node, the second node routes the data message depending on configuration data stored in its map. The map may indicate through which gateway (such as the first node) or virtual connection (e.g, a VPN including the gateway as an endpoint), if any, to forward data information to a particular (computer) destination address. The second node, such as a customer edge router at one of the remote LAN facilities, obtains map configuration information from the notification message received (e.g., using BGP) from the first node location (i.e., the first customer edge router that sent the notification message indicating the network address information and the gateway identifier). For example, transmission of data messages (in a reverse direction with respect to the notification message discussed above) may include transmitting a data message at a first computer through a CE router (e.g., second node) across a service provider network to another CE router (e.g., first node), where the message is then forwarded to a corresponding target computer based on a destination address of the data message. Thus, in the context of a physical network supporting connectivity through customer edge routers of a service provider network, a notification message including routing policy attributes (e.g., network address information and gateway identifier) dictates how data messages to a particular host or a range of hosts (as identified by network address

information) shall be forwarded through the second node and corresponding service provider network.

In one application, the notification message and routing policy attributes transmitted from the first node are disseminated to customer edge nodes (e.g., the second node) of a network supported by RFC 2547 based on a distribution protocol such as BGP (Border Gateway Protocol). The Border Gateway Protocol (BGP) is an interautonomous system routing protocol used by ISPs (Internet Service Providers) to exchange routing information in a network. Notably, a network supported by RFC 2547 does not inherently support encryption. However, corresponding maps (which are dynamically updated or modified as discussed above) in CE routers of the network may identify secured network connections (e.g., a VPN including an endpoint identified by the gateway identifier) on which to forward data information. Consequently, customers can achieve secure data transmissions between sites connected to an RFC 2547 network because security of the data is automatically managed by the corresponding CE routers.

As mentioned, a corresponding map at a network node identifies a policy for routing or forwarding data messages. By "policy", what is meant in one embodiments is that for packets destined for host computers that fall within one of the network address information ranges of a remote LAN facility (e.g. addresses within subnet lists specified in the map), the packets must be sent in a secure manner using the gateway identifier as an endpoint for a VPN tunnel that may or may not already exist (i.e., the VPN tunnel may already be established, or may need to be setup from the second node to the first node before packets falling within that address range can be sent from the second node). In the above example, if a security tunnel already exists between the second node and the first node and is the path specified by the map at the second node for forwarding a corresponding data message to the at least one computer, the data message is transmitted over the existing security tunnel. If a specified security tunnel does not exist, the second node and first node establish a security tunnel based on IPSec and IKE (Internet Key Exchange) for transmission of the data message.

In addition to transmitting a gateway identifier and network address information as discussed above, the first node may transmit additional routing policy attributes to the

second node to more particularly define a policy for routing the data messages on a corresponding virtual network connection through the gateway to the at least one host computer.

More specifically, one embodiment of the present invention is directed to a computer program product that includes a computer readable medium having instructions stored thereon for dynamically updating configuration data in a network. The instructions, when carried out by a processor of the data communication device, cause the processor to perform the steps of: i.) receiving a) network address information associated with at least one host computer, and b) a corresponding gateway identifier of a gateway in the physical network; ii.) generating a notification message including the network address information and the corresponding gateway identifier; and iii.) transmitting the notification message to a second node of the physical network enabling the second node to establish a virtual network connection between the second node and the first node on which to forward data messages to the at least one host computer based on the corresponding gateway identifier.

Another embodiment of the present invention is directed to a computer program product that includes a computer readable medium having instructions stored thereon for configuring a network to support routing of network messages. The instructions, when carried out by a processor of the data communication device, cause the processor to perform the steps of: i.) receiving a notification message from a sending node of the physical network, the notification message including network address information and a corresponding gateway identifier of a gateway of the physical network; and ii.) based on contents of the notification message, modifying a map at the receiving node to include the network address information and configuration data identifying at least part of a virtual network connection between the receiving node and the sending node on which to forward data messages through the gateway to a destination node.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other objects, features and advantages of the invention will be apparent from the following more particular description of preferred embodiments of the

invention, as illustrated in the accompanying drawings in which like reference characters refer to the same parts throughout the different views. The drawings are not necessarily to scale, emphasis instead being placed upon illustrating the principles of the invention.

FIG. 1 is a pictorial diagram of a communication system to dynamically configure edge nodes according to an embodiment of the invention.

FIG. 2 is a diagram of a configuration message to dynamically update edge nodes according to an embodiment of the invention.

FIG. 3 is a diagram of configuration data that is dynamically updated based on receipt of configuration messages according to an embodiment of the invention.

FIG. 4 is a pictorial diagram illustrating use of configuration data in a corresponding map of an edge node according to an embodiment of the invention.

FIG. 5 is a diagram of a computer system and its functional components supporting dynamic configuration according to an embodiment of the invention.

FIG. 6 is a flow chart illustrating a technique supporting dynamic configuration of edge routers according to an embodiment of the invention.

FIG. 7 is a flow chart illustrating a technique supporting dynamic configuration of edge routers according to an embodiment of the invention.

DETAILED DESCRIPTION

Use of IPSec to protect traffic between two VPN subnets requires the IPSec security gateways protecting the subnets to agree on a security policy. Many elements of the security policy may be configured once in the security gateway and the elements are independent of the topology of the VPN. There are at least two security attributes that may not be known a priori and require repetitive updates to all the security gateways as the network topology changes. The two security attributes include: i) trusted subnets (i.e., IP network address and mask) protected by a peer security gateway and ii) the security gateway's identity.

One embodiment of the invention leverages a use of the BGPv4 protocol to distribute these two security policy attributes between peers of a network. Since BGPv4 is used to propagate the subnets required for routing purposes, identifying the IP subnets

as 'trusted subnets' and associating a 'security gateway identity' with a set of 'trusted subnets' allows an IPSec security policy to be completed and distributed dynamically for each security gateway. The automated discovery of topology specific peer security policy attributes allows the protected VPN topology to evolve, change, and grow while even though there may be only a single provisioning action at the time of security gateway installation. It is possible to also include more granular IPSec security policy attributes (in addition to the two above) such as IP tunnel end-point identities, protocol identities and application port identities. The inclusion of these IPSec security policy attributes may allow a more refined security policy. Regardless of the number of security policy attributes propagated by BGPv4, the attributes are associated with the IP network address that BGPv4 is already propagating for basic VPN routing topology. Thus, there is no requirement that the IPSec security associations be used to exchange routing information such as configuration data and the routing scalability property of a network supported by [RFC2547] may be preserved.

Although the techniques described herein can be used in networking applications, and particularly to communication devices such as routers that provide connectivity to many remote devices through a network link, the techniques are also well-suited for other applications as well.

FIG. 1 is a diagram of communication system 100 supporting dynamic update of configuration information according to an embodiment of the present invention. As shown, communication system 100 includes host computers 160-1, 160-2, ..., 160-Z, 161, 162-1, 162-2, ..., 162-J, and networks (such as private local area networks) 150-1, 150-2, ..., 150-K that communicate with each other over virtual network 120. Virtual network 120 includes, at its periphery, customer edge nodes 140-1, 140-2, ..., 140-P that communicate with each other over service provider network 110. Service provider network 110, at its periphery, includes provider edge nodes 130-1, 130-2, ..., 130-N that also communicate with each other. In general, customer edge nodes 140 of virtual network 120 determine a policy for routing data messages (e.g., TCP/IP data packets) based on corresponding locally stored configuration data.

In one embodiment, virtual network 120 operates according to the connectivity model of RFC2547 and configuration data associated with a corresponding customer edge node 140 identifies a VPN on which to forward data messages when communicating with a particular node of communication system 100. For example, a node such as a computer in network 150-2 may attempt to transmit a message to host computer 162-1 associated with network 150-5. Prior to transmitting data messages such as data packets generated by the node at network 150-2, customer edge node 140-2 checks its corresponding configuration data to determine a policy for forwarding the data messages to host computer 162-1.

Configuration data at CE node 140-2 may indicate (depending how it is configured) that data messages from any node of network 150-2 to host computer 162-1 shall be encrypted based upon a security tunnel between customer edge node 140-2 and customer edge node 140-5. Thus, for a case in which a node in network 150-2 sends messages to host computer 162-1, virtual network 120 establishes a security tunnel between customer edge node 140-2 and customer edge 140-5 if such a tunnel does not already exist. The security tunnel (e.g., a VPN) may be established between gateways at corresponding customer edge nodes 140-2 and 140-5 based on IPSec and IKE. Thereafter, customer edge node 140-2 forwards the data messages through the security tunnel to customer edge node 140-5 which, in turn, forwards the data message to host computer 162-1. In a similar way, configuration data at other customer edge nodes 140 dictates how data messages shall be forwarded, if at all, through virtual network 120 to a corresponding target.

As briefly discussed, a service provider may deploy virtual network 120 based on an [RFC2547] service using a number of backbone tunneling techniques such as those described in [RFC2547], [MPLS-in-IP], or [PE-PE-IPsec]. [RFC2547] employs a hierarchical routing model providing scalable distribution of routing/forwarding attributes. IPSec encryption between customer edge nodes 140, as previously discussed, optionally relies upon an IP address partitioning and route forwarding state created by the [RFC2547] infrastructure, which can be deployed independently of a chosen type of backbone tunneling through virtual network 120. The CE-CE [IPSec] topology may

include a point-to-point relationship between CEs for data protection; however, the routing plane associated with the CE-CE topology leverages the [RFC2547] hierarchical routing model.

To couple a CE security policy associated with virtual network 120 and PE routing plane of service provider network 110, each CE node 140 ascertains through configuration, an administrator, or other means, whether it will be used as a security gateway (e.g., whether it will be an endpoint of a tunnel for communicating with nodes across virtual network 120).

After determining that a CE node 140 shall be a gateway, the corresponding CE node 140 advertises it's gateway identifier (e.g., "Security Gateway Identity") used for [IKE] and [IPSec] peer end-point termination to a corresponding PE node 130 in service provider network 110 (e.g., CE node 140-1 notifies PE node 130-1 via message 205-1) using a data distribution protocol such as [BGP-4]. The gateway identifier is associated with a 'trusted subnet' (e.g., network address information of one or multiple computers in communication with a corresponding network 150 whose communications shall be protected) represented as a prefix that the corresponding CE node 140 protects.

FIG. 2 is a diagram of message 205-1 including notification message 215, network address information 222, and gateway identifier 226 according to one embodiment of the invention. Network address information 222 identifies one or multiple network addresses such as network addresses associated with host computers 160-1, 160-2. Gateway identifier 226 identifies a node such as CE node 140-1 or other gateway in communication system 100.

One purpose of message 205-1 (or derivative thereof such as message 205-2) is to notify other CE nodes 140 in virtual network 120 of an end point of a security tunnel through which messages shall be forward when communicating with devices corresponding to nodes identified by network address information 222. For example, message 205-1 includes gateway attribute 220. For illustrative purposes, assume that gateway identifier 226 of gateway attribute 220 identifies CE node 140-1 as being a gateway (or a specific one of multiple gateway endpoints that terminate at CE node 140-

1) and associated network address information 222 identifies addresses of nodes associated with network 150-1 such as host computer 160-1 and 160-2.

Based on the receipt of gateway attribute 220 in message 205, CE node 140-5 updates its configuration data for establishing pathways such as security tunnels. More specifically, in the present example, CE node 140-5 updates its configuration data for establishing secured communications between CE node 140-5 and CE node 140-1 to properly forward messages generated by network 150-5 (or its associated nodes such as host computers 162-1, …, 162-J) to protected target devices (e.g., host computers 160-1 and 160-2) identified by network address information 222. In this way, gateway attribute 220 identifies a node (based on gateway identifier 226) through which messages shall be forwarded to corresponding communication devices identified by network address information 222.

In one application, gateway identifier 226 is an IPv4 address where the [IKE] and [IPSec] authentication and encryption services will be established. Upon receiving information of protected nodes (as identified by network address information 222) and corresponding gateway identifier 226 of the sending customer edge node 140, the corresponding PE node 130 (in the present example, PE node 130-1), in turn, advertises the trusted subnet prefixes (network address information 222) and the associated gateway identity 226 to other PE nodes 130 based on a data distribution protocol such as MP-BGP.

Thus, a PE node 130 that receives gateway attribute 220 via an [MP-BGP] message 205 (a) identifies which VPN the prefix and security gateway identifier 226 (e.g., end-point) is associated with, and (b) advertises (e.g., via message 205-2) this information to any security gateway CE nodes 130 that belong to the VPN. Identification of which VPN an update message 205 belongs may be determined by the "route target" extended community attribute as described in RFC 2547 incorporated herein by this reference.

FIG. 3 is a diagram illustrating configuration data 310-5 (such as a table or database) according to one embodiment of the invention. As shown, configuration data 310-5 (associated with corresponding CE node 140-5) includes information identifying

how to forward data messages through CE node 140-5 to a target based dynamic updating configuration data 310 in response to receiving gateway attributes 220 in corresponding message 205. For example, based on receipt of message 205 (in FIG. 2), entry #1 in configuration data 310-5 (FIG. 3) is updated to identify that messages through CE node 140-5 to host computer 160-1 and 160-2 shall be sent through a gateway at CE node 140-1. Configuration map 310-5 will be referenced only when data is sent through CE node 140-5 with host source addresses in network 150-5 (or whatever is locally configured). An IPsec policy may require both source and destination address ranges to send data over a secure tunnel.

For example, according to one embodiment, a decision to send traffic into an encrypted VPN may be based on a combination of received policy (the destination network address range received via a notification message) and the locally defined policy (e.g., a source network address range). In such an embodiment, traffic is not encrypted merely based on receipt of a network address range with a security gateway identity. Instead, a data flow that enters an encrypted VPN must match the source and destination policy. More specifically, CE node 140-2 may protects LAN_1 150-2 while CE node 2 140-5 may protect LAN_2 150-5. If CE node 1 advertises LAN_1 with protection, CE node 2 will only send data destined for LAN_1 into the encrypted VPN if it matches the combined policy of protect traffic from LAN_2 to LAN_1. Thus, one embodiment of the invention supplies the destination prefix using a scalable distribution means while the source prefix is locally defined. The source prefix may be manually configured.

Other entries of configuration data 310-5 reflect updates prompted by receipt of corresponding messages 205. For example, entry #4 reflects receipt of message 205 including a gateway identifier of host computer 160-Z and network address information 222 identifying host computer 160-Z. In this example, an identified end of a secured tunnel is host computer 160-Z. Thus, a VPN may extended beyond CE nodes for yet additional protection.

Note that each CE node 140 includes configuration data 310 that varies depending on how data messages shall be forwarded through virtual network 120. For example, CE

node 140-1 includes configuration data 310-1, CE node 140-2 includes configuration data 310-2, etc.

FIG. 4 is a diagram illustrating a process of forwarding data messages 405 through virtual network 120 according to an embodiment of the invention. As shown, CE node 140-5 receives messages 405-1, ..., 405-C. Prior to forwarding messages 405, CE node 140-5 compares destination addresses associated with messages 405 generated by network 150-5 (or associated nodes such as host computers 162) and compares them to those identified in its corresponding configuration data 310-5. Based on configuration data 310-5, CE node 140 determines a corresponding policy for forwarding messages 405. For example, entry #1 of configuration data 310-5 (FIG. 3) indicates that messages 405 to host computers 160-1 and 160-2 shall be sent on virtual network pathway 450 (such as a secured tunnel) established between CE node 140-5 and CE node 140-1 using KEY1 such as an encryption key.

In addition to specifying an endpoint, gateway or VPN on which to forward data messages, configuration data 310-5 may identify additional attributes of virtual network pathway 450 in other embodiments.

As indicated by entry # 2 of configuration data 310-5 (FIG. 3), data messages (such as any of messages 405) from network 150-5 to any of host computers 160-3, 160-4, 160-5, 160-6, and 160-7 need not be protected or sent on any particular secured pathway. For example, there is no specified gateway through which to transmit messages 405 to host computers 160-3 through 160-7.

As indicated by entry #3 of configuration data 310-5 (FIG. 3), any data messages 405 from network 150-5 (or associated nodes such as host computers 162) sent through CE node 140-5 to any network address of a node associated with network 150-2 shall be forwarded through a secured tunnel between CE node 140-5 to CE node 140-2 using encryption KEY2. In this instance, one end-point of a corresponding secured tunnel is CE node 140-5 through which data messages are transmitted while the other end-point indicated by corresponding configuration data 310-5 is CE node 140-2.

Entry # 4 of configuration data 310-5 (FIG. 3) illustrates that an endpoint of a secured tunnel may extend to a terminal such as host computer 160-Z. In this case, CE

node 140-5 forwards messages 405 destined for host computer 160-Z through a secured tunnel established between CE node 140-5 and terminal gateway is host computer 160-Z. Thus, a gateway identifier (such as an end-point of a secured tunnel as identified by gateway identifier 226) is not limited to identifying CE nodes 140 as an endpoint of a tunnel. A terminal node such as a host computer (or multiple computers identified by, for example a subnet address) may be identified by gateway identifier 226 as an end of a tunnel (or ends of tunnels).

FIG. 5 is a block diagram of CE node 140 according to an embodiment of the present invention. As shown, CE node 140 is a computerized device including interconnect 515 such as a data bus or other circuitry interconnecting memory 112, processor 113, and communication interface 560. Processor 113 may be any type of central processing unit, microprocessor, processing device, controller of other electronic circuitry capable of accessing configuration data update application 510 to execute, run, interpret, operate or otherwise perform configuration data update application 510, thus supporting dynamic updates to configuration data 310 according to embodiments of the invention as explained herein. In other words, configuration data update application 510 may be embodied as a software program that enables CE nodes 140 such as data communication devices (e.g., CE routers) to dynamically configure themselves based on receipt of gateway attributes 220 in message 205.

Memory 112 stores configuration data update application 510, and configuration data 310 associated with a corresponding CE node 140. In general, application 510 represents software code, data and/or logic instructions executed by processor 113. When executed, processor 113 creates configuration data update process 550 including message generation 520 and map generation 530 processes, which are executed at CE nodes 140 depending on whether a CE node 140 is a sending or receiving node. For example, the message generation 530 process is executed on a sending node (such as CE node 140-1) to generate messages 205-1 including gateway attribute 220 distributed through virtual network 120 to corresponding target CE nodes 140. Map generation 530 process is executed at a receiving node (such as CE node 140-5) to update its configuration data 310 when a received gateway attribute 220 pertains to the

corresponding CE node 140 as discussed. These processes are more particularly described in connection with FIGs. 6 and 7.

FIG. 6 is a diagram of flow chart 600 according to an embodiment of the invention. In general, flow chart 600 illustrates a technique of dynamically updating configuration data 310 associated with communication 100 and, more specifically, CE nodes 140 of virtual network 120.

In step 610, a given CE node 140 receives network address information 222 associated with at least one host computer and ii) a corresponding gateway identifier 226 of a gateway (e.g., a network node) in communication system 100.

In step 620, the given CE node 140 generates a notification message 205 including network address information 222 and gateway identifier 226.

In step 630, the given CE node 140 generates message 205 to a receiving PE node 130 that, in turn, distributes gateway attribute 220 to other CE nodes 140. As previously discussed, CE nodes 140 of virtual network 120 update their corresponding configuration data 310 depending on whether information within gateway attribute 220 pertains to a VPN associated with the corresponding CE node 140. This is more particularly discussed in connection with FIG. 7.

FIG. 7 is a diagram of flow chart 700 according to an embodiment of the present invention.

In step 710, a CE node 140 receives message 205 including gateway attribute 220 from a sending CE node 140 disseminating updated routing information.

In step 720, the receiving node identifies whether received message 205 includes configuration information pertinent to the receiving node (e.g., a VPN that is to be supported by the receiving node). If so, the receiving CE node 140 modifies its corresponding configuration data 310 to include network address information 222 to later establish a virtual network connection (such as a VPN) between the receiving CE node 140 and sending CE node 140 on which to forward future messages through a corresponding gateway (e.g., sending CE node 140) to a destination node such as a host computer identified by network address information 222.

Distribution of Gateway Endpoints

A CE node 140 may send encrypted and non-encrypted traffic (e.g., messages 405) through a PE node 130 for delivery to other members of its VPN. CE nodes 140 that belongs to an "encrypted" VPN build a Security Association (SA) with a remote CE node 140 that also belongs to the same VPN, and is a member of the encryption service.

When traffic that needs to be encrypted is sent from a CE node 140 that belongs to an "encrypted" VPN, the CE node 140 establishes a Security Association (SA) with the remote CE node 140 through which the destination of the incoming packet is reachable. To achieve this aim, the CE node 140 needs to discover the remote peer's trusted subnet prefix (e.g., based on network address information 222) and the associated security gateway identity (e.g., based on gateway identifier 226) of the peer, and then build the [IKE] and [IPSec] security association.

Discovery of end-point addresses may be achieved through direct [BGP-4] exchange with a PE node 130. If [BGP-4] is not established with a customer site, then a different discovery protocol may be used.

Certain [RFC2547] deployments use [BGP-4] on PE-CE links. Typically, these sessions only carry standard [BGP-4] attributes. As previously discussed, gateway attribute 220 supports dynamic update of configuration data at CE nodes 140. In one embodiment, CE nodes 140 support the capabilities as specified in [EXTCOM].

When a CE node 140 advertises routes from an "encrypted" VPN into the backbone, it attaches a new BGP extended-community attribute (e.g., gateway attribute 220) to all trusted subnet prefixes for which encryption is desired. A PE node 130 that receives such an update exports those trusted subnet prefixes along with the "Security Gateway Identify" attribute.

A PE node 130 that receives the update advertises the trusted subnet prefixes and security gateway identities to any relevant CE nodes 140 that are (a) members of the "encrypted" VPN, and (b) are running [BGP-4] with the PE node 130.

The Extended Community Attribute is a transitive optional BGP attribute, with type code 16, as specified in [EXTCOM]. One embodiment involves a use of the Opaque Extended Community, as specified in section 6.4 of [EXTCOM]. The value of the high-

order octet of this extended type may be either 0x03 or 0x43. The low-order octet of this extended type carries the sub-type with a specified value and indicates that it is a "Security Gateway Identity." The following 48 bits provide the data corresponding to the "Security Gateway Identity" (e.g., gateway identifier 226).

CE Node Supporting IPsec

A CE node 140 that wishes to belong to an "encrypted" VPN, and use the techniques herein, may depend on the procedures described in [IPsec]. For example, a CE node 140 may provide a Security Policy Database (SPD), as described in section 4.4.1 of [IPSec], which is used to determine the disposition of all IP traffic inbound or outbound from the node. Each entry within the database may specify whether traffic matching the policy should bypass IPsec processing, be discarded, or be subject to IPSec processing.

A CE node 140 may support the ability to specify "Selectors," as described in section 4.4.2 of [IPSec]. These selectors may be a set of IP and upper layer protocol field values that are used by a Security Policy Database (SPD) (e.g., configuration data 310-5) to map traffic to a Security Association (SA).

The Security Policy Database and Selector attributes may be populated with the "Security Gateway Identity" and the associated "trusted subnet" prefixes. The population of the SPD from [BGP-4] may be an automated process with the appropriate [BGP-4] controls provided by a CE node 140.

A CE node 140 may enable the creation of security associations in the Security Association Database (SAD), as described in section 4.4.3 of [IPSec], that contains parameters derived by traffic matching the [BGP-4] injected Selectors in the SPD. This database is used to determine what [IPSec] services are offered to IP packets.

CE-CE Security Association

A CE node 140 may support an automated Security Association/Key management protocol for the purpose of establishing and maintaining Security Associations between two [IPSec] peer end-points. One example of such a protocol is [IKE].

There are several options available to CE nodes 140 with respect to IPSec tunnel setup and encryption of traffic. For example, CE-CE authentication and/or encryption of selective packets may be based on traffic flow initiated establishment of security associations. Additionally, CE-CE authentication and/or encryption of selective packets may be based on pre-established [IPSec] security associations. Each of these options is more particularly described later in this specification.

Regardless of which option is used, on receipt of traffic that is matched to an SPD policy that requires [IPSec] processing, a CE node 140 checks whether a Security Association (SA) already exists with the [IPSec] Security Gateway address. If a SA (e.g., a secured tunnel) already exists, then the CE node 140 encrypts the traffic and forwards it over a secured tunnel toward a PE node 130 . If no SA exists, the CE node 140 uses [IKE] or similar protocol to establish the SA with the security gateway identity.

CE-CE Encryption of Selective Packet Flow

CE-CE encryption may be driven by traffic flow and a CE node 140 may choose to selectively encrypt packets based on a 'Selector' match. On receipt of a packet that is matched by the CE node's configuration data for encryption, the CE node 140 establishes a SA with the remote CE router through which the destination is reachable.

As a CE node 140 is running [BGP-4] with a PE node 130, it can dynamically build the 'Selector' criteria based on receipt of routing updates that carry gateway attribute 220. Using this information, the CE node 140 identifies which routes are associated with a remote site, and also which of these routes needs encryption. For the routes that need encryption, the CE determines a "Security Gateway Identity" associated with those routes.

A CE node 140 dynamically establishes [IPSec] SA's between CE and PE routers. These [IPSec] tunnels may be used to protect the [BGP-4] exchange of 'Trusted Subnets' and 'Security Gateway Identities' between the PE and CE nodes. Alternatively, CE and PE nodes may use [BGP-MD5] on the [BGP-4] session to authenticate network address information 222 and associated gateway identifier 226.

CE-CE Encryption with Pre-established IPsec Security Associations

A CE node 140 may pre-establish [IPSec] tunnels between CE nodes 140 based on configuration data 310. [IPSec] SA's may be established automatically upon population of the SPD that occurs upon receipt of a 'Trusted Subnet' prefix with a valid "Security Gateway Identity". The CE node 140 encrypts traffic destined to a route via the established [IPSec] security association.

The CE node 140 may pre-establish [IPSec] SA's between the CE and PE nodes. These [IPSec] tunnels may be used to protect the [BGP-4] exchange of 'Trusted Subnets' and 'Security Gateway Identities' between the PE and CE nodes. Alternatively, the CE and PE nodes may use [BGP-MD5] on the [BGP-4] session to authenticate the prefixes and the associated "Security Gateway Identity".

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.